Considerations for the Alternate Assessment based on Modified Achievement Standards

Table of Contents

Chapter 1: Introduction

The contents of this publication were developed under cooperative agreement S283B050019 with the U. S. Department of Education. However, the contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

For the full version of this white paper, visit:
http://nycomprehensivecenter.org/initiatives/inits_sp_assessment

University of the State of NewYork
State Education Department

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Marianne Perie

In April 2007, as part of its governance of the federal *No Child Left Behind* Act, the U.S. Department of Education (USED) released new regulations that allowed for the use of an alternate assessment based on modified achievement standards (AA-MAS). These regulations supplemented the most recent Elementary and Secondary Education Act legislation regarding the development of grade-level assessments and alternate assessments based on alternate achievement standards. States were required to develop these assessments in reading and math, grades 3–8 plus one grade in high school, and use those to hold schools and districts accountable for student progress. States could use this new assessment for students with disabilities to count up to two percent of students as Proficient for purposes of Adequate Yearly Progress (AYP). These regulations were in response to state concerns that there were students with disabilities who were not able to show proficiency on the general assessment and yet would not be assessed appropriately by the alternate assessment based on alternate achievement standards either; they fell into the "gap" between the two assessments. It supplemented the option of developing alternate assessments based on grade-level achievement standards which only a few states used.

In spring 2008, six states submitted their "modified" assessments for Peer Review to determine whether they could be used for purposes of AYP. In June 2008, USED released a paper written by Janet Filbin that describes the issues raised during the Peer Review of the six state alternate assessments based on modified achievement standards. None of the states received approval for their AA-MAS, but lessons learned from the review of their designs provided much information for all states.

In Fall 2008, the New York Comprehensive Center (NYCC) applied for and received supplemental funding from USED to collaborate with the New York State Department of Education (NYSED) to study these issues further. Their proposal involved convening national research experts to provide guidance to NYSED regarding the feasibility of developing an alternate assessment based on modified achievement standards and advice on how to design standards and assessments for students with disabilities who are part of the "2%" gap. NYCC partnered with the National Center for the Improvement of Educational Assessment (Center for Assessment) to convene a panel of experts and write a white paper on this topic.

In January 2009, a group of 17 research experts were identified and brought together in New York City to discuss the issues surrounding the design and development of an AA-MAS. This report is a result of that meeting. Nine of the experts authored chapters of this report and the remaining eight experts reviewed the chapters and provided support and guidance to the authors. Each author is a nationally recognized expert on the issue discussed in his/her chapter and the reviewers possess similar qualifications allowing for a full review both within and across chapters. A full list of the experts involved in this study can be found in Appendix A.

The purpose of this report is to describe the primary challenges in developing an AA-MAS based both on the Filbin (2008) paper and the panel's own experiences. It provides a research-based analysis of the design and development issues and focuses on the theory behind each issue. In addition, this report explores the existing research and best practices in identifying and assessing these students. Specifically, the goal of this report is to make recommendations to NYSED about developing an AA-MAS, with the expectation that these recommendations could be generalized to other states. The authors each approached their chapters with an intention to help states think through the issues, make appropriate decisions regarding the allocation of resources, and ultimately improve opportunities for students with disabilities.

Upon the completion of the second drafts of these chapters, the expert panel recognized the utility of the information beyond the application to fulfilling the federal regulations regarding an alternate assessment for purposes of accountability. Much of the discussion in this report relates to instructing and assessment all low achievers and specifically on low achievers with disabilities. Therefore, even if the regulations were rescinded, the panel believes the information provided in this report will continue to be applicable as the field works to improve our knowledge and understanding of how low achieving students with disabilities—and perhaps those without disabilities who are also struggling with grade-level achievement standards—learn, organize information, and communicate their understanding.

Background on Federal Regulations Guiding the AA-MAS

The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), known as No Child Left Behind (NCLB), required that all states assess all students in reading[1] and mathematics in grades 3 through 8 plus one grade in high school. In addition, they were required to assess all students in science at least once in elementary, middle, and high school. A minimum of three performance levels had to be developed for each test—one defining proficiency, one above that and one below that—with the goal of all students reaching proficiency by 2014. Up to 1% of students with the most significant cognitive disabilities could be categorized as proficient using an alternate assessment based on alternate achievement standards (AA-AAS). States also had the option of developing an alternate assessment based on grade-level achievement standards (AA-GLAS) for those students who were capable of performing on grade level but needed a format other than the traditional multiple-choice test to demonstrate their knowledge and skills.

[1] The law requires an assessment in reading, although some states include reading in a broader English Language Arts (ELA) assessment and use that to meeting NCLB requirements.

Some state and local leaders argued that there were still some students with disabilities who were not being well served by the assessment program because they were ineligible to take the AA-AAS and unable to access all of the content and skills assessed on either grade-level assessment. Prior to NCLB, many states used out-of-level testing to assess certain students with disabilities. For example, a student who was in grade 8 based on their age, but being instructed significantly below the 8th-grade level, might be administered a grade 6 test. NCLB ended that practice and enforced the IDEA principle that students should have access to the general curriculum by holding schools accountable for teaching students grade-level content.

The regulations released in April 2007 allowed states to develop an alternate assessment based on modified achievement standards and use it for accountability purposes. Before these regulations were released, students with disabilities had the option of taking: (1) a general grade-level assessment, with or without accommodations; (2) an alternate assessment based on grade-level achievement standards; or (3) an alternate assessment based on alternate achievement standards. Critics argued that none of these options seemed appropriate for certain groups of students with disabilities. They wanted an additional option for an appropriate assessment of what these students know and can do across all the content standards not only for accountability purposes but also to provide information that could help guide instruction. The AA-MAS was intended to fall between an AA-AAS and a general grade-level assessment to provide a more appropriate measure of these students' performance against academic content standards for the grade in which they are enrolled. The regulations state that “there is a small group of students whose disability has precluded them from achieving grade-level proficiency and whose progress is such that they will not reach grade-level proficiency in the same time frame as other students” (34 C.F.R. Part 200). However, this statement has raised countless questions as state policymakers try to determine who this “small

group" is within that larger group of students who are not eligible for the AA-AAS but who are not performing well on the grade-level assessment.

An emphasis of the regulations and the nonregulatory guidance was that this assessment must be challenging for these students. The assessments are required to cover the same breadth and depth as the other grade-level assessments. Modified achievement standards were described as being challenging for eligible students although defining a less rigorous expectation of mastery of grade-level academic content standards. They could not be linked to content from a lower grade level or exclude content standards that were assessed by the grade-level general assessment. States also were not permitted to apply their new modified achievement standards to that same general assessment; a new assessment must be developed. Students assessed using the AA-MAS must have access to grade-level content and be working towards achieving grade-level goals. However, it is important to note that the regulations do not require states to develop this assessment and provide flexibility for states to develop an AA-MAS only for a particular grade or subject.

Eligible students include students with a disability in any of the 13 disability categories defined in the Individuals with Disabilities Education Act (IDEA). To determine eligibility, the guidance stipulates:

- There must be objective evidence demonstrating the student's disability has precluded the student from achieving grade-level proficiency.
- The student's progress to date in response to appropriate instruction, including special education services designed to meet the individual needs of the student, is such that even if significant growth occurs, the IEP team is reasonably certain the student will not reach grade-level proficiency within the year covered by the IEP.
- The student's IEP must include goals that are based on the academic content standards for the grade in which the student is enrolled.

States must establish participation guidelines for IEP teams to use to match the student to the appropriate test, typically the grade-level assessment with or without accommodations, an AA-GLAS, an AA-MAS or an AA-AAS. The guidelines must include criteria based on evidence that demonstrate the student meets the three eligibility requirements bulleted above. Students should not be locked into taking an AA-MAS every year, but must have the opportunity to move from the AA-MAS to a general or alternate grade-level assessment from one year to the next. Also, a student might take the AA-MAS in one subject but the general assessment in another. All of these decisions would be made each year by the student's IEP team.

The 2007 regulations allow states to count up to 2% of students as proficient using an alternate assessment based on modified achievement standards. The number "2%" was considered to be a "reasonable cap" based on the research available to the federal government. While states were developing the AA-MAS, they were allowed to use a "2% proxy" for interim flexibility. That is, they could calculate the percentage of students with disabilities that is equivalent to 2.0 percent of all students assessed in a particular school or district. This proxy could then be added to the percentage of students with disabilities who scored proficient or above and used in making AYP determinations. This interim flexibility was first introduced in 2005 when the announcement was made that the USED was working on regulations for the AA-MAS, and it is set to expire after the 2008-09 accountability year. Using an AA-MAS, states could count up to 2% of students as proficient, replacing the 2% proxy but still providing additional flexibility for state, district, and school accountability.

Setting the Stage for this Report

A driving question for New York State (and other states) is whether the development of this assessment will yield useful information to guide instruction and be cost effective. More specifically, in times of budget cutbacks, how can the limited funding available be best allocated to support the learning of these students? The first issue in answering this question is

determining who “these students” are. Subsumed within that question is the possibility of expanding this report beyond the current federal regulations, focusing on students who may not be receiving grade-level content. In addition, it is important to consider the challenges of using the data to “guide instruction” when the primary focus of many of these assessments is simply to provide an additional measure for purposes of accountability. While more description regarding the students and the uses of the assessment will be provided more fully in the following chapters, it is important to provide a context and lay out the assumptions that this report will follow regarding fidelity to the federal regulations and guidance.

Authors were encouraged to adhere to the law laid out in the most recent IDEA and ESEA reauthorization and to stay true to the federal principles. However, if there were aspects of the April 2007 regulations permitting the development of the AA-MAS that authors found too constrictive, they were encouraged to address areas for change. Recognizing that people disagree on the assumptions behind this 2% population, this paper is written from the assumption that all students can learn (and should be taught) grade-level content standards with appropriate instruction and support. However the degree to which all students achieve the grade-level content standards may vary. Of course, even these assumptions lead to more questions about whether students are taught the exact same content or whether it will be modified as well as the time frame in which they are expected to learn the content. These more specific issues will be addressed in the first section of the report, but the authors started from these basic principles and assumptions.

Issues Specific to New York State

In elementary and middle school, the New York State Education Department (NYSED) requires NCLB testing for English and Mathematics in grades 3–8, and science assessment in grades 4 and 8. In addition to the NCLB-required tests, New York State assesses social studies

in grades 5 and 8; and technology education in grade 8. All tests are comprised of both multiple-choice and constructed-response items. Student performance is divided into four levels:

- Level I: 'Not Meeting Learning Standards'
- Level II: 'Partially Meeting Learning Standards'
- Level III: 'Meeting Learning Standards'
- Level IV: 'Meeting Learning Standards with Distinction'

New York State counts Level III and IV as Proficient for purposes of AYP.

At high school, NYSED administers Regents Examinations that are tied to the high school diploma a student receives. Currently, students are required to take Regents Examinations in English, mathematics, social studies, and science. The diploma a student receives is tied both to the courses taken and the score on the Regents Examinations. Currently, a student with the most significant cognitive disabilities typically receives an IEP diploma. Students who do not achieve the level of performance necessary to earn a Regents diploma can earn a Local diploma; a score of 65 or higher qualifies a student for a Regent's diploma; a score of 55-64 qualifies a student for a Local diploma. For purposes of AYP at the high school level, the Regents Examinations in Comprehensive English and Integrated Algebra are used. Instead of using the graduation cut scores, NYSED established three separate achievement levels for Comprehensive English and Integrated Algebra to be used solely for the purpose of calculating AYP. The Regents Examinations also are comprised of both multiple-choice and constructed-response items.

NYSED is primarily interested in exploring the AA-MAS in English and Mathematics. At this point, it has received federal approval to develop AA-MAS only in grades 3–8; however, for purposes of this report, authors were asked to consider the full range of K–12 assessments. NYSED wants to follow the regulations of assessing the same content on the AA-MAS as on its general assessment, to better understand how to make the assessment less rigorous, and to

learn how to modify the achievement standards while maintaining the reliability and validity of the results. Specific questions raised by NYSED include:

1. Which students are best served by this assessment?
2. How different are they from the rest of the special education population?
3. What is an "appropriately challenging" achievement standard?
4. Which modifications make the most sense in the context of the AA-MAS?
5. How do the modifications affect the validity and reliability of the interpretation?
6. What is the credential that is most appropriate for students participating in the AA-MAS and what does it lead to in terms of post-secondary potential?

Most of the issues raised by NYSED are general issues that many state policymakers are confronting, and many of these match closely with the issues raised by Filbin (2008). The one exception is the last question regarding student credentialing. Because the Regents examinations are used both for AYP purposes (thus open to modification) and graduation requirements, NYSED raises a good question regarding whether using an AA-MAS would limit a student's opportunity to receive a Regents diploma. The nonregulatory guidance clarifies that no assumption is made about the comparability between the AA-MAS and an assessment required for graduation. However, states may not require students to enter a non-diploma track if they take an AA-MAS. However, since the diploma in New York State is based primarily on the score obtained on a Regents exam, modifying that exam does seem to ensure the student will be tracked to a lower-level diploma. This question is a policy issue rather than a technical question, and while it will be addressed within this report, it is ultimately an issue that will need to be decided by NYSED.

Organization of the Paper

The direction to the expert panel from the Assistant Commissioner of the Office of Standards, Assessments, and Reporting in NYSED was to provide information on current

research and best practices and to make recommendations on the steps NYSED should take towards designing an AA-MAS (or to recommend not to do it at all). As a first step, the expert panelists reviewed Filbin (2008) to determine key issues. Filbin identified five areas that were challenging to states:

1. Identifying students eligible to take the AA-MAS.
2. Providing guidelines for writing standards-based IEPs and then monitoring the implementation of those guidelines.
3. Designing an assessment based on grade-level content standards that is of an appropriate difficulty and depth of knowledge for this population.
4. Determining the relationship between the AA-MAS, the general assessment, and the alternate assessment based on alternate achievement standards (AA-AAS).
5. Writing appropriate modified achievement level descriptors.

The expert panel used these five issues as a starting point during the initial planning meeting. Later, the specific questions from NYSED were added and divided among the different chapters as appropriate. Ultimately, though, this report was organized into three sections focusing on different aspects of designing and developing the AA-MAS, with three chapters in each section. Within the ten chapters (including this introduction), all of the issues described by Filbin (2008) and the questions raised by NYSED are addressed.

Section I. Identifying and Understanding the Population. The first issue raised by Filbin (2008) and asked by NYSED involves determining who should take this assessment. During the initial expert panel meeting, the experts decided that the issues of identifying the students were wrapped up in the NRC assessment triangle of assessment, instruction, and cognition (Pellegrino, Chudowsky, & Glaser, 2001). Thus it was decided that this first section should discuss the issues of identifying the students and understanding their cognitive abilities, including the interaction between instruction, cognition, and assessment. This section could be titled: who are the students, vis à vis the curriculum?

Chapter 2, written by Rachel Quenemoen, focuses on identifying students appropriate for this assessment. She provides a policy context and summarizes research related to the teaching and learning of students with disabilities. Most importantly, she lays out a framework for state policymakers in considering how to identify students who might benefit most from an alternate assessment based on modified achievement standards. Included in this framework is a discussion on improving student access to grade-level curriculum and providing more opportunities to learn.

Chapter 3, by Meagan Karvonen, takes this argument one step further by examining various instructional strategies for teaching students with disabilities, with a focus on the issue of writing standards-based IEPs. She discusses the importance of aligning the curriculum with the grade-level content standards and providing supports for students to access this curriculum within the IEPs. She describes ways to promote quality of instruction and provide guidance to IEP teams.

Finally, in chapter 4, Jim Pellegrino provides information on the third vertex of the triangle: student cognition. He discusses the importance of understanding student learning characteristics and cognitive processes in assessment, and goes on to describe possible sources of differences among students that have implications for learning, instruction, and assessment.

Section II. Test Development. This next section starts the discussion on test development. The main question the authors wrestled with was how to make the assessment more accessible for students with a wide range of disabilities but maintain the reliability and validity of the results. A deep understanding of the content and test design was necessary as well as an understanding of what is meant by modified achievement standards.

Thus, chapter 5, written by Robert Rickelman and David Pugalee, begins this section with a discussion of the content domains of reading and mathematics. They continue the discussion from the first section regarding aligning curriculum, instruction, and assessment but

focus specifically on issues related to reading and mathematics. They describe important issues regarding sampling the domain and making the content accessible to students with disabilities.

Chapter 6, by Stephen Dunbar and Catherine Welch, then moves the discussion into the issues of test development. They discuss the challenge of developing items and test forms in reading and mathematics that better match the learning characteristics of the population identified for the AA-MAS, focusing on reducing the difficulty while maintaining the reliability of the assessment.

Next, the issue of developing modified achievement level descriptors is discussed by Marianne Perie in Chapter 7. This chapter focuses on determining how the modified achievement standards "fit" between the grade-level achievement standards and the alternate achievement standards, and provides practical advice for writing achievement level descriptors and setting cut scores, discussing the theory behind each.

Section III. Technical Considerations and Practical Applications. The third section of this report addresses three issues related to the technical quality and use of the assessments: examining the validity of these assessments, determining the comparability of these assessments to the general assessment, and understanding how these assessments will be operationalized and used in a state accountability system.

In chapter 8, Jamal Abedi explores issues of comparability of the AA-MAS with the general assessments and grade-level achievement standards. The chapter is written from the premise that issues concerning comparability of assessments are of paramount importance for inclusion, as states may not produce valid outcomes if the degree of comparability across the assessments has not been clearly explored and described. Likewise, descriptions of any differences in interpretations of achievement levels of the same name across each type of assessment need to be provided. This chapter examines content and construct comparability, linguistic comparability, psychometric comparability, and the use of accommodations to achieve comparability.

Chapter 9, by Scott Marion, focuses on creating a validity argument for alternate assessments based on modified achievement standards. This chapter describes the importance of stating the policymakers' values explicitly and laying out a theory of action for the purpose and use of these assessments. It then goes on to describe types of validity evidence that can be gathered throughout the test development process and beyond and used to evaluation the assumptions in the validity argument.

Finally, in chapter 10, Chris Domaleski provides practical advice and a theoretical discussion of using these assessments in state accountability systems. The focus of this chapter is how the AA-MAS fits into state accountability systems, but specific advice is given on how to develop participation guidelines, evaluate the reliability and validity of accountability decisions made using these assessments, operationalize the "2% cap," create score reports, and use the results to determine diploma eligibility.

At the end of this white paper are three appendices followed by a glossary of terms. Appendix A simply provides information on the team that developed this white paper as the chapters were shaped by the entire expert panel. Appendix B provides suggested resources, available on the Internet, for effective curriculum and instruction. Appendix C is a tool that state policymakers can use as they are considering whether and how to develop an AA-MAS. This tool consists of questions for state policymakers and educators to consider at each phase of assessment development as well as a link back to resources within this report that will inform the discussions. Many of the questions come from the validity framework that guides much of the discussion in these chapters (c.f., Marion, 2007). Finally, a glossary of terms is included that encompasses vocabulary used in both the assessment and disabilities worlds.

References

Filbin, J. (2008). *Lessons from the initial peer review of alternate assessments based on modified achievement standards* .Washington DC: U.S. Department of Education, Office of Elementary and Secondary Education (OESE) Student Achievement and School Accountability Program.

Marion, S. (2007). *A technical design and documentation workbook for assessments based on modified achievement standards.* Minneapolis: National Center on Educational Outcomes. Retrieved April 2009 from http://cehd.umn.edu/nceo/Teleconferences/AAMASteleconferences/AAMASworkbook.pdf

Pellegrino, J. W., Chudowsky, N. J., & Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy of Sciences.

U. S. Department of Education (2007a). *Final Rule 34 CFR Parts 200 and 300: Title I—Improving the academic achievement of the disadvantaged; Individuals with disabilities education act (IDEA).* Federal Register. 72(67), Washington DC: Author. Retrieved July 12, 2008, from the World Wide Web: http://cehd.umn.edu/NCEO/2percentReg/FederalRegApril9TwoPercent.pdf

U. S. Department of Education (2007b). *Modified achievement standards: Non-regulatory guidance.* Washington DC: Office of Elementary and Secondary Education (OESE). Retrieved from the World Wide Web on August 27, 2008, at http://www.ed.gov/policy/speced/guid/nclb/twopercent.doc